

Best estimate reserves and the claims development results in consecutive calendar years

Annina Saluz*

Department of Mathematics, 8092 Zurich, Switzerland

Alois Gisler

Department of Mathematics, 8092 Zurich, Switzerland

Abstract

The claims development result (CDR) is the difference between the best estimate predictions of the ultimate claim in 2 successive years. With best estimate reserves it is often argued that CDR's in consecutive years should fluctuate randomly around zero. However, in practice one frequently observes that CDR's in a given line of business have the same sign over several consecutive years. We show that this is a phenomenon which is not unusual and to be expected in situations of change. Moreover, we show how situations of change can adequately be described by a model, taking into account the evolving external information.

Keywords

Claims reserving; Best estimate reserves; Claims development result; Diagonal effects

1. Introduction

1.1. A fundamental question

Accurate claims reserves are essential for an insurance company. If the reserves are wrong, the premiums will be wrong, which would have serious consequences on the profitability. Claims reserves are by far the most important item on the liability side of the balance sheet. They have a big impact on the profit and loss account of a given calendar year and a change of the reserves by a small percentage of say 3% might well change the yearly result from positive to negative and vice versa.

For these reasons the management wants to be confident that the actuarial reserves are best estimates, that is, reserves, which are neither optimistic nor pessimistic, neither on the prudent nor on the aggressive side. In order to calculate best estimate reserves at time t the available information needs to be taken into account. At time $t+1$ more information is available and the best estimate reserves need to be updated accordingly. The difference between two successive best estimate predictions is the claims development result (CDR). An easy and obvious first check for management is to look at the CDR's over several consecutive years. As long as the CDR's in different years fluctuate around zero, there is no reason to question that the reserves are best estimates. But as soon as the CDR in a certain line of business (lob) has the same sign over several consecutive years, be it negative or positive, the management and/or the executive directors responsible for the profitability

*Correspondence to: Annina Saluz, ETH Zurich, RiskLab, Department of Mathematics, 8092 Zurich, Switzerland.
Tel: +41 44 632 68 30. Fax: +41 44 632 15 23. E-mail: annina.saluz@math.ethz.ch

of the lob might become suspicious. If the CDR's are negative, the actuaries will be blamed to be responsible for the resulting losses in the profit and loss account. If the CDR's are positive they will be blamed that the company has missed profitable opportunities because of too high and not competitive premiums due to pessimistic reserves.

However, do CDR's with the same sign over several consecutive calendar years necessarily mean that the reserves are not best estimates? Are CDR's with the same sign over several consecutive calendar years a contradiction to best estimate reserves? This is the fundamental question we are going to discuss in this paper.

In order to answer this question we have to model situations of change in a mathematical way. A suitable modelling of realistic situations of change is a second aim of the paper. For this purpose we will consider Bayesian models with calendar year effects. For related structural ideas on the behaviour of best estimates and the updating process, we refer to Arjas (1989). In a recent paper on dependence modelling in claims run-off triangles Merz *et al.* (2013) define a log-normal model that allows to model calendar year effects. Moreover, their paper gives a summary on existing literature on calendar year effects modelling. An example for a Bayesian claims reserving model is the Bayesian model for the chain ladder (CL) method introduced by Gisler & Wüthrich (2008). Bühlmann *et al.* (2009) derive a recursive formula for the estimation of the CL factors in the model of Gisler & Wüthrich. The updating of the CL factors in this recursive procedure is based on new data. This is the usual way of Bayesian modelling: one has one prior distribution and one considers the posterior distribution, given the data. More literature on Bayesian claims reserving is, for instance, given in De Alba (2002) or Wüthrich & Merz (2008).

In contrast to many other reserving models we suggest a model that allows to update a priori assumptions due to new external information. We provide a recursive procedure with explicit formulas for the updating of the parameters. In our recursive procedure the parameter updates are not only based on new data, but also based on new external information. In other words, we consider models, in which the prior distribution can be modified dynamically. As stated in De Jong & Zehnwirth (1983) external (or subjective) information is of crucial importance in times of rapid change. De Jong & Zehnwirth (1983) use a state space model where the updating process is done via Kalman filter equations. It is not the primary goal of our paper to introduce a new reserving method or to start a discussion on different reserving methods, but rather to give an answer to the fundamental question. Nevertheless, the framework introduced in section 3.3 is of interest on its own as it allows to model several realistic situations of a changing environment. We believe that this kind of modelling could also be beneficial in pricing and other areas.

Organisation of the paper: In the remainder of this section we introduce the problem from a mathematical point of view. In section 2 we look at an example from pricing, where the basic properties to be discussed can most simply be explained. Then we will introduce in section 3.2 a simple model for reserving, where we can explicitly calculate the best estimate reserves. We provide a numerical example to illustrate the results. Finally, in section 3.3 we present a more general model that allows to model many realistic situations of a changing environment. In both models for reserving the a priori assumptions can be updated due to new external information.

1.2. Best estimate reserves

In claims reserving one usually considers a claims development triangle or trapezoid:

$$\mathcal{D}_t = \{X_{i,j} : i = 0, \dots, t; j = 0, \dots, J; i+j \leq t\}, \quad (1)$$

where $X_{i,j}$ denotes the incremental claim (incurred or paid claims) of accident year i at the end of development year j . Let $C_{i,j}$ be the corresponding cumulative claim and denote by U_i the ultimate claim amount of accident year i . If all claims are settled at the end of development year $J \leq t$ then $U_i = C_{i,J}$.

The outstanding liabilities of accident year $i \geq t - J$ at the end of calendar year t are defined by

$$R_i^{(t)} := \sum_{j=t-i+1}^J X_{i,j}^{\text{paid}} = U_i - C_{i,t-i}^{\text{paid}},$$

where $X_{i,j}^{\text{paid}}$ are the incremental claim payments of accident year i in development year j and $C_{i,t-i}^{\text{paid}}$ the cumulative claim payments of accident year i until the end of calendar year t .

For claims reserving one has to predict these outstanding liabilities, which is equivalent to predicting the ultimate claim, that is

$$\widehat{R}_i^{(t)} = \widehat{U}_i^{(t)} - C_{i,t-i}^{\text{paid}},$$

where $\widehat{U}_i^{(t)}$ is a predictor of the ultimate claim U_i at the end of calendar year t .

Best estimate reserves are reserves that are as accurate as possible and for best estimate reserves the expected value of the CDR is zero. In mathematical terms one has to be a bit more precise. Denote by \mathcal{A}_t the σ -algebra containing all information available at time t . Best estimate reserves of accident year i at time t are then defined as the reserves minimising the conditional mean square error of prediction (mse_p):

$$\text{mse}_{p_{R_i^{(t)}}|\mathcal{A}_t}(\widehat{R}_i^{(t)}) = E\left[\left(\widehat{R}_i^{(t)} - R_i^{(t)}\right)^2 \middle| \mathcal{A}_t\right]$$

and are given by

$$\widehat{R}_i^{(t)} = E\left[R_i^{(t)} \middle| \mathcal{A}_t\right] = E[U_i | \mathcal{A}_t] - C_{i,t-i}^{\text{paid}}. \quad (2)$$

The best estimate reserve at the end of calendar year t is then the sum over all accident years:

$$\widehat{R}^{(t)} = \sum_{i=t-J+1}^t \widehat{R}_i^{(t)}.$$

Remarks.

- The available information \mathcal{A}_t might consist of two components $\mathcal{A}_t = \sigma(\mathcal{D}_t, \mathcal{I}_t)$ where \mathcal{D}_t are the observed data until time t and where \mathcal{I}_t denotes some additional external information, independent of the data such as observations and expert opinions with regard to legal and economic developments. In the actuarial literature often only the information contained in the data \mathcal{D}_t is considered. However, we believe that additional available information should also be taken into account. In mathematical terms the sequences $\{\mathcal{A}_t : t = 0, 1, \dots\}$ as well as $\{\mathcal{D}_t : t = 0, 1, \dots\}$ and $\{\mathcal{I}_t : t = 0, 1, \dots\}$ are filtrations.
- $E[U_i | \mathcal{A}_t]$ is not exactly known at time t , as it depends itself on unknown parameters. For instance, in the CL model of Mack (1993) $E[U_i | \mathcal{A}_t]$ depends on the unknown CL factors f_j , which have to be estimated from the data. Thus, to be strict, the best estimate reserve is only an estimate of the conditional expected value of the outstanding liabilities. However, we would like to study a model where best estimate reserves can explicitly be calculated in order to show that CDR's with the same sign over several consecutive calendar years are not a contradiction to best estimate reserves. This means that even if we know the parameters exactly, CDR's with the same sign over several

consecutive calendar years are a phenomenon that is to be expected in certain situations. Therefore, it will be assumed that the best estimate reserves as defined in equation (2) are exactly known.

1.3. The martingale argument

The CDR of accident year $t-J \leq i \leq t-1$ in the financial statement of calendar year t is defined by

$$\text{CDR}_i^{(t)} := \widehat{R}_i^{(t-1)} - X_{i,t-i}^{\text{paid}} - \widehat{R}_i^{(t)}, \quad (3)$$

where $X_{i,t-i}^{\text{paid}}$ denotes the incremental claim payments of accident year i in calendar year t . This can also be written as

$$\text{CDR}_i^{(t)} = \widehat{U}_i^{(t-1)} - \widehat{U}_i^{(t)}.$$

Since

$$\widehat{U}_i^{(t)} = E[U_i | \mathcal{A}_t],$$

the sequence

$$\left\{ \widehat{U}_i^{(t)} : t = i, i+1, \dots \right\} \quad (4)$$

is a martingale with respect to the filtration $(\mathcal{A}_t)_{t \in \mathbb{N}}$. Hence the sequence

$$\left\{ \text{CDR}_i^{(t)} : t = i+1, i+2, \dots \right\},$$

representing CDR's in consecutive years, consists of increments of a martingale. Therefore, these CDR's are uncorrelated and have expected value zero. From this fact it is often concluded that CDR's having the same sign over several consecutive calendar years are a strong indication that the reserves are not best estimates.

Remark.

Also with a martingale it is not necessarily unlikely that the increments have the same sign over several consecutive years. This is, in particular, the case if the distribution of the CDR's is strongly skewed to the left and if the probability of a positive CDR is much higher than 0.5. This can be the case in special lob such as product liability with potential large incurred but not reported claims occurring with a small probability. But even if the distribution is symmetric around zero and if we assume that CDR's of different calendar years are independent, within a long-time horizon of say 50 years there is, for instance, a probability of 54.4% to have a consecutive period of 6 or more years with the same sign of the CDR. However, such long-time horizons are not considered in insurance, and for a reasonable time horizon of say 10 years this probability reduces to 9.4%. Thus, the martingale argument is not a very strict argument from a mathematical point of view but it reinforces the general belief, that CDR's with the same sign over several consecutive years are a contradiction to best estimate reserves.

1.4. Observations in practice

In practice, however, one often observes the same sign of the CDR over several consecutive years. A striking example was the claims development after the mid-1990s in the line motor liability in Switzerland. Most of the Swiss insurance companies experienced a negative CDR over a longer period of about 8 consecutive years. The reason was the emergence of a new phenomenon called "whiplash". It is mainly observed after rear-end collisions. The victims complain about permanent pain, chronic fatigue, loss of concentration and inability to work. However, as a peculiarity of this phenomenon, there is no objective medical injury that could be diagnosed by a medical doctor and

that would explain these symptoms. Because of this lack of an observable medical indication the courts did not acknowledge a disability for such cases for a long time. But after a decision at the Swiss Federal court, the court practice began gradually to change. This was the beginning of a long process leading to a permanently increasing claims load, which was impossible to predict and that manifested itself only with time. For instance, there are many rear-end collisions, but fortunately most of them turned out to be harmless and only a small percentage developed to a case with long-term injury. However, at the beginning, there was no claims experience allowing to estimate neither this percentage nor the expected claim amount of the corresponding claims. There was also a change in the whole environment. For instance, more and more lawyers and law offices began to specialise on whiplash cases which itself increased the number of persons claiming for a disability. Or the advocates of the victims were not prepared to settle a claim in the hope that time works in favour of the victims and that they can get a higher amount at a later settlement date. It was also a cultural issue, as the whiplash phenomenon was mainly observed in the German speaking part of Switzerland and was practically non-existent in the French and Italian speaking part. The question was whether or when it would spill over to these parts of Switzerland, too. Fortunately it did not.

What are best reserve estimates in such periods of change and what should a reserving actuary do? We claim that it is not sufficient just to look at the claims triangle. In the paid claims triangle there is usually hardly seen any change at the early years of such a period and the new situation is reflected in the claims data only gradually and with a delay. In the incurred triangles one might observe an increase in the newer diagonals. However, such a new situation is also a challenge for the claims adjusters and there is a big uncertainty in their case estimates and as a consequence in the information observed in the newer diagonals of the incurred triangle. One of the authors has experienced other situations where the claims adjusters overreacted at the beginning. We claim that best practice is to communicate closely with the claims department, to take their judgement and evaluation as an expert opinion, which is permanently updated based on the development at the claims front (verdicts, court practice, etc.) and on the observed data in the claims triangles. In the motor liability example of the mid-1990s it took several years until the situation stabilised and the new court practice was reflected in the claims data, and hence the resulting CDR's were negative over several years. In 2008, there was another decision at the Swiss Federal court with the effect that courts have become much more reluctant to attribute a permanent disability after an accident with whiplash indications. After that there was observed a similar development of the CDR's in the opposite direction.

The Swiss motor liability example shows that also with best reserving practice the CDR's can have the same sign over a longer period. The question arises whether we have overseen something in the martingale argument? The answer is yes, we have overseen something.

There is nothing wrong in the mathematics and the sequence $\{\widehat{U}_i^{(t)} : t = i, i+1, \dots\}$, as defined in equation (4), is a martingale. However, one has to be careful in the interpretation what this martingale property exactly means and what conclusions one can draw from the martingale argument. The point is that the claim payments and the claims development depend on “hidden”, not directly observable, characteristics describing the “state space” at different times such as legal and economic environment. These “hidden” characteristics are best modelled in a Bayesian way as random variables. The martingale argument only holds in the average over these state space characteristics. But what we observe in a claims development triangle are conditional observations, that is, claims data generated on the condition that a specific but unknown realisation of the state space variables is given. The observed CDR's resulting from best estimates based on such conditional observations are no longer increments of a martingale with expected value zero. This fact will become more clear when looking at some concrete examples.

2. An example from pricing

The Poisson-Gamma model is a well-known model for experience rating based on observed claim frequencies and was for instance the basis for constructing the Bonus-Malus scheme in motor liability in Switzerland (see Bichsel, 1964).

Assume we have a portfolio of risks numbered $i = 1, 2, \dots, I$ and denote by $N_{i,j}$ the claim number of risk i in year j . Each risk is characterised by its risk characteristic ϑ_i which is a realisation of a random variable Θ_i .

Model Assumptions 2.1 (Poisson-Gamma)

- i) Conditionally, given $\Theta_i = \vartheta_i$, the random variables $N_{i,j}$ ($j = 1, 2, \dots$) are independent and Poisson distributed with Poisson parameter $\vartheta_i \lambda_0$.
- ii) The random variables Θ_i are independent and Gamma distributed with $E[\Theta_i] = 1$ and $\text{Var}(\Theta_i) = \gamma^{-1}$.
- iii) $N_{i,j}$ and $N_{k,l}$ are independent for $i \neq k$.

Denote by $\mathcal{D}_t = \{N_{i,j} : j = 1, 2, \dots, t; i = 1, 2, \dots, I\}$ the observations up to year t . Based on these observed data we want to predict for each risk i next years claim number. The following theorem is well-known from credibility literature (see, for instance, Bühlmann & Gisler, 2005, Chapter 2).

Theorem 2.2 *The best prediction of $N_{i,t+1}$ given \mathcal{D}_t under Model Assumptions 2.1 is given by*

$$\hat{N}_{i,t+1} = \alpha_t \bar{N}_{i,t} + (1 - \alpha_t) \lambda_0,$$

where

$$\bar{N}_{i,t} = \frac{1}{t} \sum_{j=1}^t N_{i,j} \quad \text{and} \quad \alpha_t = \frac{t\lambda_0}{t\lambda_0 + \gamma}.$$

Remarks.

- The individual risk characteristics ϑ_i are drawn at the very beginning from a first “urn” as realisations of independent and identically distributed Gamma random variables.
- The sequence $\{\hat{N}_{i,t} : t = 1, 2, \dots\}$ is a martingale with respect to $(\mathcal{D}_t)_{t \in \mathbb{N}}$ and it holds that

$$E[\hat{N}_{i,t+1} - N_{i,t+1} | N_{i,1}, \dots, N_{i,t}] = 0. \quad (5)$$

This means that for an infinitely large portfolio there is a balance between forecasted and future claim frequency in all sub-portfolios of drivers having the same claim experience $N_{i,1}, \dots, N_{i,t}$. This is the best we can achieve for a risk-adjusted pricing. However, the left-hand side of equation (5) is an expected value over the whole Θ space.

- But the martingale property does not hold for the observations of a specific driver i . A specific driver i has its specific risk characteristics $\Theta_i = \vartheta_i$ drawn from a first urn, and the observed claim numbers $N_{i,j}$ of this particular driver are independent and Poisson distributed with Poisson parameter $\vartheta_i \lambda_0$. We then have

$$\begin{aligned} E[\hat{N}_{i,t+1} - N_{i,t+1} | N_{i,1}, \dots, N_{i,t}, \Theta_i] &= \hat{N}_{i,t+1} - E[N_{i,t+1} | \Theta_i] \\ &= \alpha_t (\bar{N}_{i,t} - \lambda_0) + \lambda_0 (1 - \Theta_i), \end{aligned}$$

which is in general different from zero. Moreover,

$$\begin{aligned} E\left[\widehat{N}_{i,t+1} - N_{i,t+1} \mid \Theta_i\right] &= \alpha_t(\lambda_0\Theta_i - \lambda_0) + \lambda_0(1 - \Theta_i) \\ &= \lambda_0(1 - \alpha_t)(1 - \Theta_i). \end{aligned}$$

Thus, our best estimate predictor has conditional on Θ_i an upward ($\Theta_i < 1$) or downward ($\Theta_i > 1$) bias for all years t . This bias decreases when the observed number t of years increases.

In reserving we do not have a portfolio of different similar claims triangles and we are not interested in the mean over these triangles. We are rather looking at one specific triangle with given hidden state space characteristics comparable to a specific driver in the pricing example. Conditional on this hidden risk characteristics the best estimate reserves are no longer unbiased, and this bias can have the same sign over several consecutive calendar years. If the 1-year process variance is small compared to this bias, we will observe a CDR with the same sign over several consecutive years. In section 3 we will consider examples from reserving.

3. Examples from reserving

3.1. Basic Bayesian framework

It is our personal belief, based on many years of practical experience of one of the authors, that reserving is as much art as science. By this we mean that a good reserving actuary in practice should not only analyse claims data from a claims triangle. There is much more information available which one could and should take into account. The lawyers and claims adjusters in the company will have some knowledge about changes in legislation, about claims handling, claims processes and the assessment of case estimates by the claims adjusters. Economists will have some opinion about the future development of economic factors. In short, there is a lot of expert knowledge available in a company from which one can learn something about the expected future claim payments. Beside the data in the triangle this expert knowledge should also be taken into account when calculating best estimate claims reserves.

In mathematical terms such a situation can best be modelled in a Bayesian framework, which is predestinated to deal with a priori knowledge on the one hand and observed claims data on the other. In such a framework the best estimate reserve is given by equation (2). Often there is no explicit formula for equation (2) and in this case a good approximation could be to use credibility theory instead. There is one difference to the credibility models usually considered in the literature: the a priori information does also change in time. In suitable models this can still be handled by recursive estimators.

The basic Bayesian framework considered in this paper is a multiplicative model with the Bayesian structure given below in Model Assumptions 3.1. Recall that \mathcal{I}_t denotes the external information at time t and \mathcal{D}_t is defined as in equation (1). By calendar year effect we mean that

$$E[X_{i,j} \mid \Psi, \mathcal{I}_t] = E[X_{i,j} \mid \Psi_{i+j}] = \mu_i \gamma_j \Psi_{i+j}, \quad t \geq 0.$$

Since the Ψ_t are multiplicative factors to be applied on the conditional expected value of all elements of a diagonal, they are also called diagonal effects. We denote by $\Psi = \{\Psi_0, \Psi_1, \dots\}$ the set of diagonal effects in all calendar years $t \in N$.

Model Assumptions 3.1 (Bayes Model with Diagonal Effects) Conditionally, given Ψ and \mathcal{I}_t for any $t \geq 0$, the increments $X_{i,j}$ are independent and there exist constants μ_i , $0 \leq i \leq t$, and γ_j , $0 \leq j \leq J$, with $\sum_j \gamma_j = 1$ such that

$$E[X_{i,j} | \Psi, \mathcal{I}_t] = \mu_i \gamma_j \Psi_{i+j}.$$

Here $\{\gamma_j, j = 0, \dots, J\}$ is the incremental payment pattern and $\{\beta_j, j = 0, \dots, J\}$ with $\beta_j = \sum_{k=0}^j \gamma_k$ the cumulative payment pattern.

Remark.

If $\Psi_t = 1$ for all $t \geq 0$, then there are no calendar year effects and if we omit the external information \mathcal{I}_t in Model Assumptions 3.1 we obtain the basic assumptions of the Bornhuetter–Ferguson (BF) model (see, for instance, Wüthrich & Merz, 2008). The best estimate reserve for accident year i at time t is then given by

$$\hat{R}_i = (1 - \beta_{t-i}) \mu_i$$

which is the BF reserve formula.

Under Model Assumptions 3.1 the best estimate reserve of accident year i at the end of calendar year t is given by

$$\begin{aligned} \hat{R}_i^{(t)} &= E[R_i | \mathcal{I}_t, \mathcal{D}_t] = E[E[R_i | \mathcal{I}_t, \mathcal{D}_t, \Psi] | \mathcal{I}_t, \mathcal{D}_t] \\ &= \mu_i \sum_{j=t-i+1}^J \gamma_j E[\Psi_{i+j} | \mathcal{I}_t, \mathcal{D}_t], \end{aligned} \quad (6)$$

and the best estimate of the total reserve at the end of calendar year t is

$$\begin{aligned} \hat{R}^{(t)} &= \sum_{i=t-J+1}^t \hat{R}_i^{(t)} \\ &= \sum_{i=t-J+1}^t \mu_i \sum_{j=t-i+1}^J \gamma_j E[\Psi_{i+j} | \mathcal{I}_t, \mathcal{D}_t] \\ &= \sum_{k=1}^J \left(E[\Psi_{t+k} | \mathcal{I}_t, \mathcal{D}_t] \sum_{j=k}^J \gamma_j \mu_{t+k-j} \right). \end{aligned} \quad (7)$$

Hence to find the best estimate reserve $\hat{R}^{(t)}$ we have to determine

$$\Psi_{t+k}^{(t)} = E[\Psi_{t+k} | \mathcal{I}_t, \mathcal{D}_t]$$

for $k = 1, \dots, J$. Note that $\Psi_{t+k}^{(t)}$ is the best k -year prediction of Ψ_{t+k} at time t .

One year later there are the observations of the diagonal of the next calendar year $t+1$ and possibly updated external information \mathcal{I}_{t+1} available. With these new informations we get the updated forecasts of $\Psi_{t+k}^{(t+1)}$.

From equation (7) it follows that the total CDR in calendar year $t+1$ is given by

$$\begin{aligned}
 \text{CDR}_{t+1} &= E \left[\sum_{i=0}^t U_i \middle| \mathcal{I}_t, \mathcal{D}_t \right] - E \left[\sum_{i=0}^t U_i \middle| \mathcal{I}_{t+1}, \mathcal{D}_{t+1} \right] \\
 &= \sum_{i=t-J+1}^t \mu_i \sum_{j=t-i+1}^J \gamma_j \Psi_{i+j}^{(t)} - \sum_{i=t-J+1}^t X_{i,t-i+1} - \sum_{i=t-J+2}^t \mu_i \sum_{j=t-i+2}^J \gamma_j \Psi_{i+j}^{(t+1)} \\
 &= \sum_{k=t+1}^{t+J} \Psi_k^{(t)} \sum_{j=k-t}^J \mu_{k-j} \gamma_j - \sum_{j=1}^J X_{t+1-j,j} - \sum_{k=t+2}^{t+J} \Psi_k^{(t+1)} \sum_{j=k-t}^J \mu_{k-j} \gamma_j \\
 &= \sum_{l=1}^J \Psi_{t+l}^{(t)} \sum_{j=l}^J \mu_{t+l-j} \gamma_j - \sum_{j=1}^J X_{t+1-j,j} - \sum_{l=2}^J \Psi_{t+l}^{(t+1)} \sum_{j=l}^J \mu_{t+l-j} \gamma_j. \quad (8)
 \end{aligned}$$

In the special case where $\mu_i = \mu$ for all i , equations (7) and (8) simplify to

$$\widehat{R}^{(t)} = \mu \sum_{k=1}^J \Psi_{t+k}^{(t)} (1 - \beta_{k-1}),$$

$$\begin{aligned}
 \text{CDR}_{t+1} &= \underbrace{\Psi_{t+1}^{(t+1)} \mu (1 - \beta_0) - \sum_{j=1}^J X_{t+1-j,j}}_{\text{difference between updated forecast and observations}} + \underbrace{\sum_{k=1}^J \left(\Psi_{t+k}^{(t)} - \Psi_{t+k}^{(t+1)} \right) \mu (1 - \beta_{k-1})}_{\text{change in update}} \quad (9)
 \end{aligned}$$

$$\begin{aligned}
 &= \underbrace{\Psi_{t+1}^{(t)} \mu (1 - \beta_0) - \sum_{j=1}^J X_{t+1-j,j}}_{\text{difference between forecast and observations}} + \underbrace{\sum_{k=2}^J \left(\Psi_{t+k}^{(t)} - \Psi_{t+k}^{(t+1)} \right) \mu (1 - \beta_{k-1})}_{\text{change future}}. \quad (10)
 \end{aligned}$$

In the following we consider two specific models within this basic Bayesian framework in which we can explicitly calculate best estimate reserves.

3.2. One Jump Model

We consider the simple situation with a stable period in the years $t \leq s_0 - 1$ with no diagonal effects observable in the claims triangles and with no indication that this situation will change in the future. However, in the year s_0 , there is a change in the environment (e.g. legislation), which leads to a change of the expected value of the total claim payments in each “cell” $\{(i, j): i+j \geq s_0\}$ by some unknown factor Ψ_{s_0} . There is no indication that further changes will incur after year s_0 . After calendar year s_0 the situation is again stable. This situation is reflected by the following model.

Model Assumptions 3.2 (One Jump Model)

i) Conditionally, given Ψ and \mathcal{I}_t for any $t \geq 0$, the $X_{i,j}$ are independent and normally distributed with

$$E[X_{i,j} | \Psi, \mathcal{I}_t] = \mu_i \gamma_j \Psi_{i+j}, \quad \text{Var}(X_{i,j} | \Psi, \mathcal{I}_t) = \mu_i \eta_j^2 \sigma^2,$$

where $\sum_{j=0}^J \gamma_j = \sum_{j=0}^J \eta_j^2 = 1$.

ii) For $t \leq s_0 - 1$, conditionally on \mathcal{I}_t and \mathcal{D}_t , $\Psi_l = 1$ for $l \leq t$ and

$$E[\Psi_{t+l} | \mathcal{I}_t, \mathcal{D}_t] = 1, \quad l \geq 1.$$

iii) For $t \geq s_0$, conditionally, given \mathcal{I}_t and \mathcal{D}_{t-1}

$$\Psi_l = 1, \quad l \leq s_0 - 1,$$

$$\Psi_l = \Psi_{s_0}, \quad s_0 \leq l \leq t,$$

$$E[\Psi_{t+l} | \mathcal{I}_t, \mathcal{D}_t] = E[\Psi_{s_0} | \mathcal{I}_t, \mathcal{D}_t], \quad l \geq 0,$$

where $\Psi_{s_0} | (\mathcal{I}_t, \mathcal{D}_{s_0-1}) \stackrel{d}{=} \Psi_{s_0} | (\mathcal{I}_{s_0}, \mathcal{D}_{s_0-1}) \sim N(m, \tau^2)$.

Remarks.

- The stable situation without diagonal effect in the period $t \leq s_0 - 1$ and no indication for a change in the future is modelled by ii) ($\Psi_l = 1$ for $l \leq t$ and prediction $E[\Psi_{t+l} | \mathcal{I}_t, \mathcal{D}_t] = 1$ for future years).
- The change of future payments by an unknown factor due to the change of the environment in the year s_0 is modelled by $\Psi_{s_0} | (\mathcal{I}_{s_0}, \mathcal{D}_{s_0-1}) \sim N(m, \tau^2)$ in iii). Note that this means a change of the “expert opinion” with regard to the years $t \geq s_0$.
- The stable situation with no further changes after the year s_0 is modelled by $\Psi_{s_0} | (\mathcal{I}_t, \mathcal{D}_{s_0-1}) \stackrel{d}{=} \Psi_{s_0} | (\mathcal{I}_{s_0}, \mathcal{D}_{s_0-1})$ and the fact that there is no indication for a change in the future by $E[\Psi_{t+l} | \mathcal{I}_t, \mathcal{D}_t] = E[\Psi_{s_0} | \mathcal{I}_t, \mathcal{D}_t]$ for $l \geq 0$.

In order to simplify the formulas we assume in the following that $s_0 \geq J$.

Proposition 3.3 Under Model Assumptions 3.2 Ψ_{s_0} is conditioned on \mathcal{D}_{s_0+l} and \mathcal{I}_{s_0+l} , $l \geq 0$, normally distributed with

$$\Psi_{s_0}^{(s_0+l)} = E[\Psi_{s_0} | \mathcal{I}_{s_0+l}, \mathcal{D}_{s_0+l}] = \alpha_{s_0+l} \bar{D}_{s_0+l} + (1 - \alpha_{s_0+l})m,$$

$$q_{s_0}^{(s_0+l)} = \text{Var}(\Psi_{s_0} | \mathcal{I}_{s_0+l}, \mathcal{D}_{s_0+l}) = (1 - \alpha_{s_0+l})\tau^2,$$

where

$$\bar{D}_{s_0+l} = \frac{\sum_{t=s_0}^{s_0+l} \sum_{j=0}^J X_{t-j,j} \frac{Y_j}{\eta_j^2}}{\sum_{t=s_0}^{s_0+l} \sum_{j=0}^J \frac{\mu_{t-j} \gamma_j^2}{\eta_j^2}} \quad \text{and} \quad \alpha_{s_0+l} = \frac{\sum_{t=s_0}^{s_0+l} \sum_{j=0}^J \frac{\mu_{t-j} \gamma_j^2}{\eta_j^2}}{\sum_{t=s_0}^{s_0+l} \sum_{j=0}^J \frac{\mu_{t-j} \gamma_j^2}{\eta_j^2} + \frac{\sigma^2}{\tau^2}}.$$

Remark.

Proposition 3.3 still yields the credibility estimator if we drop the normality assumption for Ψ_{s_0} and the $X_{i,j}$. In a distribution-free approach, it is the best linear approximation to the Bayes estimator and in this sense still a best estimate. However, it is then no longer exact Bayesian and a best estimate as defined in equation (2).

Proof of Proposition 3.3:

In the One Jump Model 3.2 we have exact credibility, that is, the credibility estimator is exact Bayesian (see, e.g. Jewell, 1974a, 1974b or Bühlmann & Gisler, 2005). If $\gamma_j \neq 0$ for $j = 0, \dots, J$, then the random variables

$$Y_{i,j} = \frac{X_{i,j}}{\mu_i \gamma_j}, \quad i+j \geq s_0,$$

fulfil the assumptions of the Bühlmann–Straub model with weights $w_{i,j} = \mu_i \gamma_j^2 / (\sigma^2 \eta_j^2)$. The formula then follows from the Bühlmann–Straub credibility formula. For incurred claims some of the γ_j might be zero. The formula can then be derived with the regression credibility model (see Theorem 8.4 in Bühlmann & Gisler, 2005) by using the observation vector $X^{(s_0+l)}$ with all observations from calendar year s_0 to calendar year s_0+l as elements, where the regression equation is given by $E[X_{i,j} | \Psi_{s_0}] = \mu_i \gamma_j \Psi_{s_0}$ and where the conditional covariance matrix of $X^{(s_0+l)}$ given Ψ_{s_0} is diagonal with diagonal elements $\text{Var}(X_{i,j} | \Psi_{s_0}) = \mu_i \eta_j^2 \sigma^2$.

The predictors $\Psi_{s_0}^{(t)}$ and the variances $q_{s_0}^{(t)}$ can also be calculated recursively (see Theorem 9.6 in Bühlmann & Gisler, 2005):

Proposition 3.4 *We have the following recursive formulas:*

i) For $t \leq s_0 - 1$

$$\Psi_{s_0}^{(t)} = 1.$$

ii) For $t = s_0$

$$\Psi_{s_0}^{(s_0)} = \frac{\sum_{j=0}^J X_{s_0-j,j} \frac{\gamma_j}{\eta_j^2}}{w_{s_0}} + (1 - \tilde{\alpha}_{s_0})m, \quad q_{s_0}^{(s_0)} = (1 - \tilde{\alpha}_{s_0})\tau^2,$$

where

$$\tilde{\alpha}_{s_0} = \frac{w_{s_0}}{w_{s_0} + \frac{\sigma^2}{\tau^2}} \quad \text{and} \quad w_{s_0} = \sum_{j=0}^J \mu_{s_0-j} \frac{\gamma_j^2}{\eta_j^2}.$$

iii) For $t > s_0$

$$\Psi_{s_0}^{(t)} = \frac{\sum_{j=0}^J X_{t-j,j} \frac{\gamma_j}{\eta_j^2}}{w_t} + (1 - \tilde{\alpha}_t) \Psi_{s_0}^{(t-1)}, \quad q_{s_0}^{(t)} = (1 - \tilde{\alpha}_t) q_{s_0}^{(t-1)},$$

where

$$\tilde{\alpha}_t = \frac{w_t}{w_t + \frac{\sigma^2}{q_{s_0}^{(t-1)}}} \quad \text{and} \quad w_t = \sum_{j=0}^J \mu_{t-j} \frac{\gamma_j^2}{\eta_j^2}.$$

With Proposition 3.4 we can also find an explicit formula for the CDR. For simplicity we consider the case with $\mu_i \equiv \mu$. By using equations (9) and (10) we obtain:

• for $t+1 \leq s_0-1$

$$\begin{aligned} \text{CDR}_{t+1} &= \mu(1 - \beta_0) - \sum_{j=1}^J X_{t+1-j,j} \\ &= \text{difference between forecast and observations,} \end{aligned}$$

- for $t+1 = s_0$

$$\begin{aligned} \text{CDR}_{s_0} &= \underbrace{\Psi_{s_0}^{(s_0)} \mu(1-\beta_0) - \sum_{j=1}^J X_{s_0-j,j}}_{\text{difference between updated forecast and observations}} + \underbrace{\left(1 - \Psi_{s_0}^{(s_0)}\right) \mu \sum_{k=1}^J (1-\beta_{k-1})}_{\text{change in update}} \\ &= \underbrace{\mu(1-\beta_0) - \sum_{j=1}^J X_{s_0-j,j}}_{\text{difference between forecast and observations}} + \underbrace{\left(1 - \Psi_{s_0}^{(s_0)}\right) \mu \sum_{k=2}^J (1-\beta_{k-1})}_{\text{change future}}, \end{aligned}$$

and

$$1 - \Psi_{s_0}^{(s_0)} = 1 - m + \tilde{\alpha}_{s_0} \left(m - \frac{\sum_{j=0}^J X_{s_0-j,j} \frac{\gamma_j}{\eta_j^2}}{w_{s_0}} \right),$$

- for $t+1 > s_0$

$$\begin{aligned} \text{CDR}_{t+1} &= \underbrace{\Psi_{s_0}^{(t+1)} \mu(1-\beta_0) - \sum_{j=1}^J X_{t+1-j,j}}_{\text{difference between updated forecast and observations}} + \underbrace{\left(\Psi_{s_0}^{(t)} - \Psi_{s_0}^{(t+1)}\right) \mu \sum_{k=1}^J (1-\beta_{k-1})}_{\text{change in update}} \\ &= \underbrace{\Psi_{s_0}^{(t)} \mu(1-\beta_0) - \sum_{j=1}^J X_{t+1-j,j}}_{\text{difference between forecast and observations}} + \underbrace{\left(\Psi_{s_0}^{(t)} - \Psi_{s_0}^{(t+1)}\right) \mu \sum_{k=2}^J (1-\beta_{k-1})}_{\text{change future}}, \end{aligned}$$

and

$$\Psi_{s_0}^{(t)} - \Psi_{s_0}^{(t+1)} = \tilde{\alpha}_{t+1} \left(\Psi_{s_0}^{(t)} - \frac{\sum_{j=0}^J X_{t+1-j,j} \frac{\gamma_j}{\eta_j^2}}{w_{t+1}} \right).$$

We observe that in the case $t+1 > s_0$ the second terms “change in update” and “change future” respectively, converge to 0 as $t \rightarrow \infty$ since the weights $\tilde{\alpha}_{t+1} \rightarrow 0$. On the other hand, the difference between (updated) forecast and observations depends more on the volatility of the data.

Next, we want to consider the sign of the CDR over several consecutive calendar years in the period starting at year $s_0 + 1$. What we will observe are conditional data generated by the model, given a specific realisation of Ψ_{s_0} . For this reason we first investigate the behaviour of the conditional expected value $E[\Psi_{s_0}^{(t)} | \mathcal{I}_{s_0}, \mathcal{D}_{s_0-1}, \Psi_{s_0}]$.

With the estimator $\Psi_{s_0}^{(t)}$ in Proposition 3.3 and from $E[\bar{D}_t | \mathcal{I}_{s_0}, \mathcal{D}_{s_0-1}, \Psi_{s_0}] = \Psi_{s_0}$, $t \geq s_0$, it follows that

$$E[\Psi_{s_0}^{(t)} - \Psi_{s_0}^{(t+1)} | \mathcal{I}_{s_0}, \mathcal{D}_{s_0-1}, \Psi_{s_0}] = (\alpha_t - \alpha_{t+1})(\Psi_{s_0} - m).$$

We then obtain for the conditional expectation of the CDR for $t + 1 > s_0$:

$$\begin{aligned}
 E[\text{CDR}_{t+1} | \mathcal{I}_{s_0}, \mathcal{D}_{s_0-1}, \Psi_{s_0}] \\
 &= \mu(1-\beta_0)(1-\alpha_t)(m-\Psi_{s_0}) + E\left[\Psi_{s_0}^{(t)} - \Psi_{s_0}^{(t+1)} | \mathcal{I}_{s_0}, \mathcal{D}_{s_0-1}, \Psi_{s_0}\right] \mu \sum_{k=2}^J (1-\beta_{k-1}) \\
 &= \mu(1-\beta_0)(1-\alpha_t)(m-\Psi_{s_0}) + (\alpha_{t+1}-\alpha_t)(m-\Psi_{s_0}) \mu \sum_{k=2}^J (1-\beta_{k-1}) \\
 &= (m-\Psi_{s_0}) \mu \left((1-\beta_0)(1-\alpha_t) + (\alpha_{t+1}-\alpha_t) \sum_{k=2}^J (1-\beta_{k-1}) \right). \quad (11)
 \end{aligned}$$

If $\beta_j \leq 1$ for all j , or $\gamma_j \geq 0$ for all j , respectively, then the second factor in equation (11) is positive since α_t is increasing in t . Moreover, this second factor decreases in t since $1-\alpha_t$ and $\alpha_{t+1}-\alpha_t$ are decreasing. In the limit for $t \rightarrow \infty$, it vanishes. In addition, we can see from equation (11) that the absolute value of the conditional expectation of the CDR is larger if more weight is given to the tail, that is, for long-tailed lob the bias is higher.

Hence, we have found the following result with regard to the behaviour of the conditional expected value of the CDR in the case where $\gamma_j \geq 0$ for all j :

- The sign of $E[\text{CDR}_{t+1} | \mathcal{I}_{s_0}, \mathcal{D}_{s_0-1}, \Psi_{s_0}]$ is the same for all $t+1 \geq s_0+1$.
- The absolute value of $E[\text{CDR}_{t+1} | \mathcal{I}_{s_0}, \mathcal{D}_{s_0-1}, \Psi_{s_0}]$ is the bigger, the bigger $|m-\Psi_{s_0}|$ is.
- The absolute value of $E[\text{CDR}_{t+1} | \mathcal{I}_{s_0}, \mathcal{D}_{s_0-1}, \Psi_{s_0}]$ is the bigger, the longer the development is. In other words, the absolute value of $E[\text{CDR}_{t+1} | \mathcal{I}_{s_0}, \mathcal{D}_{s_0-1}, \Psi_{s_0}]$ is bigger, if the development pattern $(\beta_j)_{0 \leq j \leq J}$ increases slower to $\beta_J = 1$.
- The absolute value of $E[\text{CDR}_{t+1} | \mathcal{I}_{s_0}, \mathcal{D}_{s_0-1}, \Psi_{s_0}]$ is decreasing in t .
- In the limit for $t \rightarrow \infty$ the CDR is conditionally unbiased.

3.2.1. Numerical example

For the example we choose a special case of the above model in which $\mu_i \equiv \mu$ and $\eta_j^2 = \gamma_j$. In this case the formula for $\Psi_{s_0}^{(t)}$ simplifies to

$$\Psi_{s_0}^{(t)} = \alpha_t \frac{1}{\mu(t-s_0+1)} \sum_{l=s_0}^t \sum_{j=0}^J X_{l-j,j} + (1-\alpha_t)m, \quad t \geq s_0,$$

where

$$\alpha_t = \frac{\mu(t-s_0+1)}{\mu(t-s_0+1) + \frac{\sigma^2}{\tau^2}}.$$

We choose the following parameters

$$\mu = 10,000, \quad \sigma = 4, \quad m = 1.02, \quad \tau = 0.02,$$

and the development pattern γ_j given in Table 1.

In our simulation of the specific observed triangle we start with a 10×10 triangle and we assume the jump occurs in calendar year $s_0 = 10$. We then add calendar years and calculate the CDR for time $s_0 = 10$ to time $t = 21$. The true diagonal effect to be considered as a specific realisation of

Table 1. Development pattern γ_j .

j	0	1	2	3	4	5	6	7	8	9
γ_j	0.2	0.23	0.12	0.11	0.09	0.07	0.058	0.05	0.043	0.029

Table 2. Incremental claims.

	0	1	2	3	4	5	6	7	8	9
0	1,930	2,426	1,208	1,266	903	946	636	480	389	169
1	2,185	2,372	1,099	1,110	932	793	516	452	462	362
2	2,406	2,334	1,304	959	819	696	525	466	550	294
3	2,035	2,548	1,272	1,159	1,149	871	459	324	473	375
4	2,199	2,559	1,437	1,152	1,038	657	681	575	290	283
5	2,169	2,486	1,237	951	1,052	660	692	537	496	307
6	2,164	2,277	1,564	1,035	1,098	822	459	462	597	322
7	1,918	2,495	958	1,289	862	825	663	648	484	332
8	2,347	2,186	1,267	1,123	954	892	526	492	534	281
9	2,366	2,614	1,251	1,052	961	952	597	596	607	384
10	2,138	2,581	1,408	1,291	990	758	541	581	532	308
11	1,958	2,564	1,182	1,119	739	762	609	503	347	412
12	2,068	2,540	1,389	1,181	989	895	605	552	527	343
13	1,752	2,627	1,267	1,378	788	847	517	560	337	
14	1,674	2,355	1,326	1,142	1,064	497	543	420		
15	2,245	2,176	1,172	994	1,143	770	456			
16	1,980	2,425	1,285	1,067	854	880				
17	1,992	2,420	1,420	1,020	1,010					
18	2,024	2,387	1,229	1,521						
19	2,488	2,687	1,361							
20	1,979	2,498								
21	1,816									

Ψ_{s_0} is given by $\psi_{s_0} = 1.05$. Hence, we consider a rather extreme situation where Ψ_{s_0} deviates substantially from the a priori expert's expectation of $m = 1.02$. Of course, if the realisation ψ_{s_0} coincides with m , then there is no systematic bias in the CDR_t for $t \geq s_0 + 1$, and the bigger the deviation of ψ_{s_0} from m , the bigger is this bias. The simulated incremental claims data is given in Table 2.

In Table 3 the CDR's are given with the splits in the terms “change future”, “difference between forecast and observations”, “difference between updated forecast and observations” and “change in update”. Moreover, we give the conditional expected value of the CDR, which is always negative in this example. In the CDR we observe negative values in 5 consecutive years. More precisely, we have 10 times a loss in the 12 years from calendar years 10–21. This is also illustrated in the plots in Figure 1 and Figure 2. From the table and from the figures one can see that in this specific example the deviation of the forecasts from the observations in the newest diagonal contributes much more to the observed CDR than the effect of the parameter update for future years.

The predictions $\Psi_{s_0}^{(t)}$ are given in Table 3 and a corresponding plot is given in Figure 3.

Table 3. CDR, change future (Future), difference between forecast and observations (Forecast), difference between updated forecast and observations (Up forecast), change in update (Update), conditional expected value of the CDR and predictions $\Psi_{s_0}^{(t)}$.

t	CDR_t	Future	Forecast	Up forecast	Update	$E[\text{CDR}_t \mathcal{I}_{s_0}, \mathcal{D}_{s_0-1}, \Psi_{s_0}]$	$\Psi_{s_0}^{(t)}$
10	-1,563	-718	-845	-560	-1,003	-923	1.036
11	-479	-92	-387	-351	-128	-273	1.040
12	-181	-38	-143	-128	-53	-218	1.042
13	-459	-26	-433	-423	-36	-180	1.043
14	-897	-88	-808	-773	-124	-154	1.048
15	194	8	187	184	11	-134	1.047
16	-281	-26	-256	-245	-36	-118	1.049
17	-112	-1	-111	-111	-1	-106	1.049
18	-270	-26	-244	-233	-37	-95	1.050
19	129	-32	162	175	-45	-87	1.052
20	-198	-9	-189	-186	-12	-80	1.052
21	-425	-15	-410	-404	-21	-74	1.053

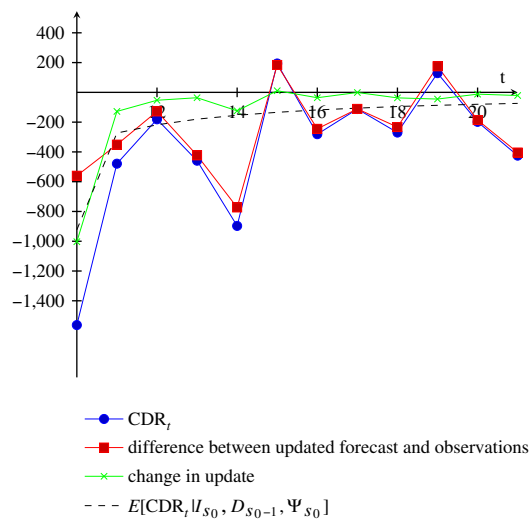


Figure 1. Claims development result (CDR) and split in “difference between updated forecast and observations” plus “change in update” and $E[\text{CDR}_t | \mathcal{I}_{s_0}, \mathcal{D}_{s_0-1}, \Psi_{s_0}]$.

For illustrative purposes we have only considered one realisation of a claims triangle with the given diagonal effect ψ_{s_0} in the above example. We additionally simulate 10,000 triangles with this diagonal effect and consider the conditional distribution of the CDR. It is important to keep the diagonal effect fixed, because in the situation considered we have exactly one effect and therefore we should not average over different effects.

We give boxplots of the realised values of the CDR in the simulations in Figure 4. The conditional expected value is negative and we observe that it gets closer to zero in later calendar years. In the first few calendar years after the jump, it is likely to observe negative CDR's.

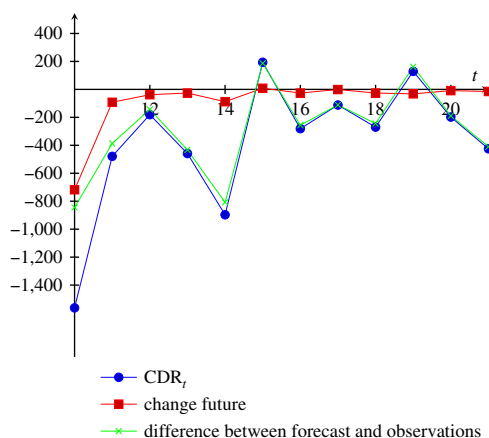


Figure 2. Claims development result (CDR) and split in “difference between forecast and observations” plus “change future”.

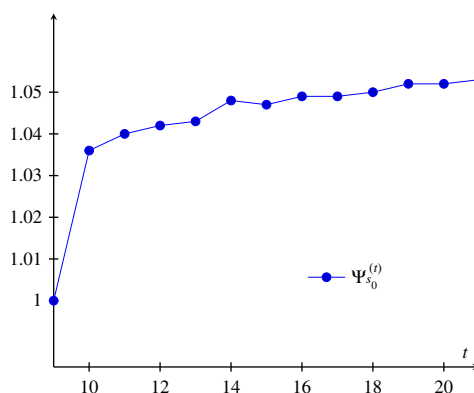


Figure 3. Predictions $\Psi_{s_0}^{(t)}$ of Ψ_{s_0} .

In practice, there are usually several consecutive jumps, often going into the same direction over a period of years. Such situations reinforce the fact of having the same sign of the CDR over several consecutive years. In the next section we give a more general framework in order to model such cases more adequately.

3.3. Normal–Normal Model

In the real world, the process with regard to change of legislation is longer and not as simple as in the One Jump Model 3.2. For instance, there can be a first pending process on a lower court, then a first verdict followed by other processes and other verdicts possibly on the next higher courts, and sometimes the legal question is even brought to the highest Swiss Federal court whose verdicts have a precedent-setting character. All this can take several years. In the mathematical set-up this means that one has several consecutive jumps going in the same direction and thus reinforcing the fact of having the same sign of the CDR over several consecutive years.

Moreover, in the real world there can be several sources and events (change of legislation, economic factors and so on) that lead to a diagonal effect. Some of them lead to an additive change (jump) in

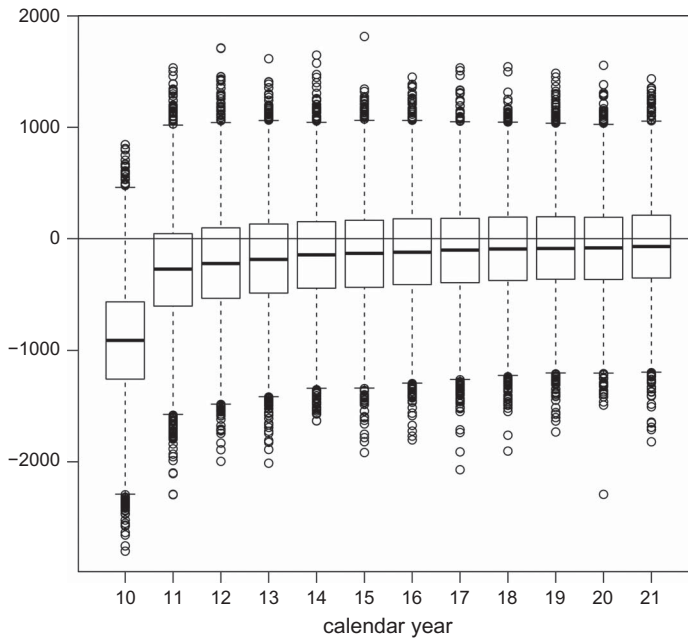


Figure 4. Boxplot of CDR_t , with 10,000 simulations, $\psi_{s_0} = 1.05$.

the diagonal effects, others to a multiplicative one or to a combination of both. For instance, the impact of economic factors such as consumer price inflation on Ψ_t could possibly be modelled by an autoregressive process meaning that

$$\Psi_t = \rho\Psi_{t-1} + (1-\rho) + \Delta_t,$$

where the Δ_t are i.i.d. with $E[\Delta_t] = 0$ (see also Shi *et al.*, 2012).

The following model is quite general and allows to mathematically describe situations, where the diagonal effect Ψ_t can change dynamically in every calendar year in an additive or multiplicative manner or both. By dynamically we mean that the parameters $a_{t+l}^{(t)}$ (multiplicative) and $b_{t+l}^{(t)}$ (additive) in Model Assumptions 3.5 depend on the information at time t and are therefore adjusted over time.

Model Assumptions 3.5 (Normal–Normal Model)

i) Conditionally, given Ψ and \mathcal{I}_t , the $X_{i,j}$ are independent and normally distributed with

$$\begin{aligned} E[X_{i,j}|\Psi, \mathcal{I}_t] &= E[X_{i,j}|\Psi_{i+j}] = \mu_i \gamma_j \Psi_{i+j}, \\ \text{Var}(X_{i,j}|\Psi, \mathcal{I}_t) &= \mu_i \eta_j^2 \sigma^2, \end{aligned} \quad (12)$$

where $\sum_{j=0}^J \gamma_j = \sum_{j=0}^J \eta_j^2 = 1$.

ii) Conditionally, given \mathcal{I}_t and \mathcal{D}_{t-1} , Ψ_t is normally distributed with

$$\Psi_t^{(t,t-1)} := E[\Psi_t | \mathcal{I}_t, \mathcal{D}_{t-1}] = a_t^{(t)} \Psi_{t-1}^{(t-1)} + b_t^{(t)},$$

$$q_t^{(t,t-1)} := \text{Var}(\Psi_t | \mathcal{I}_t, \mathcal{D}_{t-1}) = \left(a_t^{(t)}\right)^2 q_{t-1}^{(t-1)} + \delta_t^{2(t)},$$

for $t \geq 0$ and $\Psi_{-1}^{(-1)} = 1$ and $q_{-1}^{(-1)} = \tau^2$.

iii) Conditionally, given \mathcal{I}_t and \mathcal{D}_t ,

$$\Psi_{t+l}^{(t)} := E[\Psi_{t+l} | \mathcal{I}_t, \mathcal{D}_t] = a_{t+l}^{(t)} \Psi_{t+l-1}^{(t)} + b_{t+l}^{(t)}, \quad l \geq 1,$$

and $a_{t+l}^{(t)} \geq 0$ and $b_{t+l}^{(t)}$ are $\sigma(\mathcal{I}_t, \mathcal{D}_{t-1})$ measurable for $l \geq 0$.

iv) For $0 \leq l \leq t$, $\Psi_{t-l} | (\mathcal{I}_t, \mathcal{D}_{t-l})$ has the same distribution as $\Psi_{t-l} | (\mathcal{I}_{t-l}, \mathcal{D}_{t-l})$.

Remarks.

- The $a_{t+l}^{(t)}$, $b_{t+l}^{(t)}$ and $\delta_{t+l}^{2(t)}$ steering the change from Ψ_t to Ψ_{t+l} also depend on \mathcal{I}_t . Consider, for instance, the situation where there is a change of legislation in some year t and one expects that the incremental claim payments $X_{i,j}$ in future calendar years $i+j > t$ will increase by some percentage, but one is uncertain about the exact value of this percentage. Such situations can be taken into account by a suitable choice of $b_{t+l}^{(t)}$ and $\delta_{t+l}^{2(t)}$.
- Note that by assumption iii), at time t we include the possibility of expected jumps and/or trends in the future.
- In the incurred triangle, there are several possible situations such as the introduction of new guidelines in the claims department, staff fluctuation, changes in legislation, which can lead to higher uncertainties in the case estimates. Such situations can be taken into account by replacing the variance assumption (12) in the model by

$$\text{Var}(X_{i,j} | \Psi, \mathcal{I}_t) = \mu_i \eta_j^2 \sigma^2 c_{i+j}^{2(t)}, \quad (13)$$

where $c_{i+j}^{2(t)}$ are constants depending on the external information \mathcal{I}_t and $c_{i+j}^{2(t)} = c_{i+j}^{2(i+j)}$ for $i+j \leq t$.

- Special cases:

- If we assume that the diagonal effects are given by

$$\Psi_t = 1 + \Delta_t, \quad E[\Delta_t | \mathcal{I}_t] = 0,$$

with independent Δ_t then the diagonal effects of different calendar years are independent. This can be modelled by choosing $a_{t+l}^{(t)} \equiv 0$ and $b_{t+l}^{(t)} \equiv 1$ for $l \geq 0$ and all t .

- The stationary AR(1) process

$$\Psi_{t+l} = \rho \Psi_{t+l-1} + (1-\rho) + \Delta_{t+l}, \quad (14)$$

can be modelled by choosing $a_{t+l}^{(t)} \equiv \rho$, $b_{t+l}^{(t)} \equiv (1-\rho)$ and $\tau^2 = \delta^2/(1-\rho^2)$. The formula for $\Psi_t^{(t,t-1)}$ is then a weighted average between the last prediction and the initial estimate 1:

$$\Psi_t^{(t,t-1)} = \rho \Psi_{t-1}^{(t-1)} + (1-\rho).$$

Similarly for the quadratic loss $q^{(t,t-1)}$ we have the weighted average

$$q^{(t,t-1)} = \rho^2 q_{t-1}^{(t-1)} + (1-\rho^2) \tau^2.$$

- In practice one also needs to estimate the parameters γ_j and μ_i . But as already mentioned in section 1.2 this is not an essential point for the message of this paper and we therefore assume that they are known.

In the following theorem we show that under Model Assumptions 3.5 there exists a recursive procedure to calculate the predictors $\Psi_t^{(t)}$ and the corresponding conditional mseps:

$$q_t^{(t)} = \text{Var}(\Psi_t \mid \mathcal{I}_t, \mathcal{D}_t) = E \left[\left(\Psi_t - \Psi_t^{(t)} \right)^2 \mid \mathcal{I}_t \right].$$

Theorem 3.6 (recursive procedure, Normal–Normal Model) *Under Model Assumptions 3.5 $\Psi_{t+k}^{(t)}$ and $q_t^{(t)}$ can be calculated by the following recursive procedure:*

i) *Initialisation:*

$$\Psi_{-1}^{(-1)} = 1, \quad q_{-1}^{(-1)} = \tau^2.$$

ii) *For $t \geq 0$:*

a)

Update in the newest diagonal:

$$\Psi_t^{(t)} = \alpha_t \bar{X}_t + (1 - \alpha_t) \Psi_t^{(t-1)},$$

where

$$\alpha_t = \frac{w_t}{w_t + \kappa_t}, \quad \text{with} \quad w_t = \sum_{j=0}^{\min(J,t)} \mu_{t-j} \frac{\gamma_j^2}{\eta_j^2}, \quad \kappa_t = \frac{\sigma^2}{q_{(t,t-1)}^{(t)}},$$

$$\bar{X}_t = \frac{1}{w_t} \sum_{j=0}^{\min(J,t)} \frac{\gamma_j}{\eta_j^2} X_{t-j,j}$$

and

$$q_t^{(t)} = (1 - \alpha_t) q_t^{(t-1)}.$$

b)

Forecast of future diagonal effects: for $k \geq 1$:

$$\Psi_{t+k}^{(t)} = a_{t+k}^{(t)} \Psi_{t+k-1}^{(t)} + b_{t+k}^{(t)}.$$

Remarks.

- If we drop the normal assumptions and do not make any distributional assumptions, Theorem 3.6 still yields the credibility estimators.
- Special cases:
 - For independent diagonal effects $\Psi_t = 1 + \Delta_t$, with $\Delta_t \mid \mathcal{I}_t \sim N(0, \delta_t^{2(t)})$, we get the predictors

$$\Psi_t^{(t)} = \alpha_t \bar{X}_t + (1 - \alpha_t), \quad \text{where} \quad \alpha_t = \frac{w_t}{w_t + \frac{\sigma^2}{\delta_t^{2(t)}}}.$$

As in this case the diagonal effects of different calendar years are independent, $\Psi_t^{(t)}$ depends only on the observations of calendar year t .

- In case of the AR(1) process (14) we obtain

$$\begin{aligned}\Psi_t^{(t)} &= \alpha_t \bar{X}_t + (1 - \alpha_t) \left(\rho \Psi_{t-1}^{(t-1)} + (1 - \rho) \right), \\ q_t^{(t)} &= (1 - \alpha_t) \left(\rho^2 q_{t-1}^{(t-1)} + (1 - \rho^2) \epsilon^2 \right).\end{aligned}$$

Proof of Theorem 3.6: The formulas for $\Psi_t^{(t)}$ and $q_t^{(t)}$ follow from the posterior distribution of Ψ_t , given \mathcal{I}_t and \mathcal{D}_t :

$$\begin{aligned}f_{\Psi_t | \mathcal{I}_t, \mathcal{D}_t} &\propto \exp \left(- \frac{(\Psi_t - \Psi_t^{(t,t-1)})^2}{2q_t^{(t,t-1)}} \right) \prod_{i+j=t} \exp \left(- \frac{(X_{i,j} - \mu_i \gamma_j \Psi_t)^2}{2\mu_i \eta_j^2 \sigma^2} \right) \\ &\propto \exp \left(- \frac{1}{2} \left(\Psi_t^2 \left(\sum_{i+j=t} \frac{\mu_i \gamma_j^2}{\eta_j^2 \sigma^2} + \frac{1}{q_t^{(t,t-1)}} \right) - 2\Psi_t \left(\sum_{i+j=t} \frac{\gamma_j X_{i,j}}{\eta_j^2 \sigma^2} + \frac{\Psi_t^{(t,t-1)}}{q_t^{(t,t-1)}} \right) \right) \right).\end{aligned}$$

Hence we obtain

$$\Psi_t^{(t)} = E[\Psi_t | \mathcal{I}_t, \mathcal{D}_t] = \frac{\sum_{i+j=t} \frac{\gamma_j X_{i,j}}{\eta_j^2 \sigma^2} + \frac{\Psi_t^{(t,t-1)}}{q_t^{(t,t-1)}}}{\sum_{i+j=t} \frac{\mu_i \gamma_j^2}{\eta_j^2 \sigma^2} + \frac{1}{q_t^{(t,t-1)}}} = \alpha_t \bar{X}_t + (1 - \alpha_t) \Psi_t^{(t,t-1)}$$

and

$$q_t^{(t)} = \text{Var}(\Psi_t | \mathcal{I}_t, \mathcal{D}_t) = \frac{1}{\sum_{i+j=t} \frac{\mu_i \gamma_j^2}{\eta_j^2 \sigma^2} + \frac{1}{q_t^{(t,t-1)}}} = (1 - \alpha_t) q_t^{(t,t-1)}.$$

As in section 3.2 one can derive a formula for the conditional expected value of the CDR. For the CDR we obtain under Model Assumptions 3.5 from equation (8)

$$\text{CDR}_{t+1} = \Psi_{t+1}^{(t+1)} \sum_{j=1}^J \mu_{t+1-j} \gamma_j - \sum_{j=1}^J X_{t+1-j,j} + \sum_{k=1}^J \left(\Psi_{t+k}^{(t)} - \Psi_{t+k}^{(t+1)} \right) \sum_{j=k}^J \mu_{t+k-j} \gamma_j.$$

In the appendix we derive the following formula for the conditional expected value of the CDR

$$\begin{aligned}E[\text{CDR}_{t+1} | \mathcal{I}_{t+1}, \mathcal{D}_t, \Psi] &= \left(\Psi_{t+1}^{(t+1,t)} - \Psi_{t+1} \right) \left(\sum_{j=1}^J \mu_{t+1-j} \gamma_j + \alpha_{t+1} \sum_{k=2}^J \prod_{l=2}^k a_{t+l}^{(t+1)} \sum_{j=k}^J \mu_{t+k-j} \gamma_j \right) \\ &\quad + \sum_{k=1}^J \left(\Psi_t^{(t)} \left(\prod_{l=1}^k a_{t+l}^{(t)} - \prod_{l=1}^k a_{t+l}^{(t+1)} \right) + \sum_{l=1}^k \left(b_{t+l}^{(t)} \prod_{n=l+1}^k a_{t+n}^{(t)} - b_{t+l}^{(t+1)} \prod_{n=l+1}^k a_{t+n}^{(t+1)} \right) \right) \\ &\quad \times \sum_{j=k}^J \mu_{t+k-j} \gamma_j,\end{aligned}$$

where $\Psi = \{\Psi_0, \Psi_1, \dots\}$. The cases of interest are the situations where the diagonal effects change in the same direction (increase or decrease) over several years. It is likely that in such situations the effects are underestimated (as in the example of motor liability in Switzerland) and in such a case the conditional expected value of the CDR has the same sign over several consecutive years. More details on this are given in the appendix.

4. Concluding remarks

The primary aim of this paper was to show that the same sign of the CDR over several consecutive calendar years is not a contradiction to best estimate reserves. This has already been achieved by considering the One Jump Model 3.2.

More realistic situations can be mathematically well described by the Normal–Normal Model 3.5 introduced in section 3.3. This model is quite general and covers a variety of real-life situations.

Impacts of changes in the claims' environment are usually first detected in the incurred claims figures, whereas they often only appear in the paid triangle with a time lag. The reason is that the claims' adjusters can react much quicker to such changes by adjusting their case estimates accordingly, which is then seen as a diagonal effect in the incurred triangle. For this reason reserving actuaries are usually looking at both, the paid and the incurred triangle. It might happen that the claims adjusters overreact in their case estimates. However, it will not be possible to see, whether and to which extent an observed trend in the newest diagonal of the incurred triangle is systematic or due to "overreaction". But the actuary will know that there is a higher uncertainty in the incurred claims figures, which can be taken into account by the time-dependent parameters $c_{i+j}^{2(t)}$ as given in equation (13). For the best estimate reserves it turns out from Theorem 3.6 by replacing σ^2 by $\sigma^2 c_i^{2(t)}$, that the higher the uncertainty, the less weight is given to the observations in the newest diagonal. This result reflects what we expect and what is also done in practice, although more by intuition than in a systematic way.

In our model we have assumed that there is the same diagonal factor for all elements of a diagonal. There can be situations, where this not the case, for instance, if a change of legislation only has an impact on bodily injury claims. Our modelling framework can be extended to such situations by assuming that the diagonal effects are random vectors rather than random variables. But this is beyond the aim and the scope of this paper.

In this paper we have considered full Bayesian models in order to be able to explicitly calculate best estimate reserves defined as the posterior expected values of future claim payments. One might object that a normal distribution is not a realistic assumption for claims data but one can easily drop the distributional assumptions and switch to a corresponding credibility model and credibility estimators. But this is again out of the primary goal and scope of this paper. The same holds for μ_i and γ_j , which were assumed to be known. Of course, in practice one will have to estimate them.

We want to emphasise once more that in our model the expert opinion changes over time. This idea can also be found in De Jong & Zehnwirth (1983). However, in most of the actuarial literature on Bayes- and credibility models there is only one a priori distribution not changing over time. We believe that models that allow for changing expert opinions over time have a big potential to bring models nearer to the processes in the real world and to solve problems in practice, not only for reserving but also for pricing and other applications.

Acknowledgements

The authors would like to thank two anonymous referees for valuable comments that helped to improve an earlier version of this paper.

References

- Arjas, E. (1989). The claims reserving problem in non-life insurance: some structural ideas. *ASTIN Bulletin*, **19**, 139–152.
- Bichsel, F. (1964). Erfahrungstarifizierung in der Motorfahrzeug-Haftpflicht-Versicherung. *Bulletin of Swiss Association of Actuaries*, **64**(1), 119–130.
- Bühlmann, H., De Felice, M., Gisler, A., Moriconi, M. & Wüthrich, M.V. (2009). Recursive credibility formula for chain ladder factors and the claims development result. *ASTIN Bulletin*, **39**(1), 275–306.
- Bühlmann, H. & Gisler, A. (2005). *A Course in Credibility Theory and its Applications*. Springer-Verlag, Berlin.
- De Alba, E. (2002). Bayesian estimation of outstanding claims reserves. *North American Actuarial Journal*, **6**(4), 1–20.
- De Jong, P. & Zehnwirth, B. (1983). Claims reserving, state-space models and the Kalman filter. *Journal of the Institute of Actuaries*, **110**, 157–181.
- Gisler, A. & Wüthrich, M.V. (2008). Credibility for the chain ladder reserving method. *ASTIN Bulletin*, **38**(2), 565–600.
- Jewell, W.S. (1974a). Credibility means are exact Bayesian for exponential families. *ASTIN Bulletin*, **8**(1), 77–90.
- Jewell, W.S. (1974b). Exact multidimensional credibility. *Bulletin of Swiss Association of Actuaries*, **74**(2), 193–214.
- Mack, T. (1993). Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin*, **23**(2), 213–225.
- Merz, M., Wüthrich, M.V. & Hashorva, E. (2013). Dependence modelling in multivariate claims run-off triangles. *Annals of Actuarial Science*, **7**(1), 3–25.
- Shi, P., Basu, S. & Meyers, G.G. (2012). A Bayesian log-normal model for multivariate loss reserving. *North American Actuarial Journal*, **16**(1), 29–51.
- Wüthrich, M.V. & Merz, M. (2008). *Stochastic Claims Reserving Methods in Insurance*. Wiley, Chichester.

Appendix

As in the case of the One Jump Model 3.2, we want to study the signs of consecutive CDR's in the Normal-Normal Model 3.5. For this purpose we consider a conditional expected value of the CDR. Recall the CDR under Model Assumptions 3.5:

$$\text{CDR}_{t+1} = \Psi_{t+1}^{(t+1)} \sum_{j=1}^J \mu_{t+1-j} \gamma_j - \sum_{j=1}^J X_{t+1-j,j} + \sum_{k=1}^J \left(\Psi_{t+k}^{(t)} - \Psi_{t+k}^{(t+1)} \right) \sum_{j=k}^J \mu_{t+k-j} \gamma_j$$

and conditionally on $\Psi = \{\Psi_0, \Psi_1, \dots\}$ we get

$$\begin{aligned} E[\text{CDR}_{t+1} | \mathcal{I}_{t+1}, \mathcal{D}_t, \Psi] &= (1 - \alpha_{t+1}) \left(\Psi_{t+1}^{(t+1,t)} - \Psi_{t+1} \right) \sum_{j=1}^J \mu_{t+1-j} \gamma_j \\ &\quad + \sum_{k=1}^J \left(\Psi_{t+k}^{(t)} - E \left[\Psi_{t+k}^{(t+1)} \middle| \mathcal{I}_{t+1}, \mathcal{D}_t, Y \right] \right) \sum_{j=k}^J \mu_{t+k-j} \gamma_j. \end{aligned}$$

Note that for $k \geq 1$

$$\begin{aligned}\Psi_{t+k}^{(t)} &= \Psi_t^{(t)} \prod_{l=t+1}^{t+k} a_l^{(t)} + \sum_{l=t+1}^{t+k} b_l^{(t)} \prod_{n=l+1}^{t+k} a_n^{(t)}, \\ \Psi_{t+k}^{(t+1)} &= \Psi_{t+1}^{(t+1)} \prod_{l=t+2}^{t+k} a_l^{(t+1)} + \sum_{l=t+2}^{t+k} b_l^{(t+1)} \prod_{n=l+1}^{t+k} a_n^{(t+1)},\end{aligned}$$

and hence

$$\begin{aligned}\Psi_{t+k}^{(t)} - E\left[\Psi_{t+k}^{(t+1)} \middle| \mathcal{I}_{t+1}, \mathcal{D}_t, \Psi\right] \\ = \Psi_t^{(t)} \left(\prod_{l=t+1}^{t+k} a_l^{(t)} - \prod_{l=t+1}^{t+k} a_l^{(t+1)} \right) + \sum_{l=t+1}^{t+k} \left(b_l^{(t)} \prod_{n=l+1}^{t+k} a_n^{(t)} - b_l^{(t+1)} \prod_{n=l+1}^{t+k} a_n^{(t+1)} \right) \\ + \alpha_{t+1} \left(\Psi_{t+1}^{(t+1,t)} - \Psi_{t+1} \right) \prod_{l=t+2}^{t+k} a_l^{(t+1)}.\end{aligned}$$

It follows that

$$\begin{aligned}E[CDR_{t+1} | \mathcal{I}_{t+1}, \mathcal{D}_t, \Psi] &= \left(\Psi_{t+1}^{(t+1,t)} - \Psi_{t+1} \right) \left(\sum_{j=1}^J \mu_{t+1-j} \gamma_j + \alpha_{t+1} \sum_{k=2}^J \prod_{l=2}^k a_{t+l}^{(t+1)} \sum_{j=k}^J \mu_{t+k-j} \gamma_j \right) \\ &+ \sum_{k=1}^J \left(\Psi_t^{(t)} \left(\prod_{l=1}^k a_{t+l}^{(t)} - \prod_{l=1}^k a_{t+l}^{(t+1)} \right) + \sum_{l=1}^k \left(b_{t+l}^{(t)} \prod_{n=l+1}^k a_{t+n}^{(t)} - b_{t+l}^{(t+1)} \prod_{n=l+1}^k a_{t+n}^{(t+1)} \right) \right) \\ &\times \sum_{j=k}^J \mu_{t+k-j} \gamma_j.\end{aligned}$$

In order to analyse this term, we assume that $\gamma_j \geq 0$ for all j .

Let us consider the situation where a new phenomenon leads to an increase in the diagonal effects that is, $\Psi_t \leq \Psi_{t+1} \leq \dots \leq \Psi_{t+n}$ for some $n > 1$. It is likely that this increase is at first underestimated (as in the example of motor liability in Switzerland) and hence it is plausible that for $l \geq 1$ and for some $k \geq 1$:

$$\begin{aligned}a_{t+l}^{(t)} &\leq a_{t+l}^{(t+1)} \leq a_{t+l}^{(t+2)} \leq \dots \leq a_{t+l}^{(t+k)}, \\ b_{t+l}^{(t)} &\leq b_{t+l}^{(t+1)} \leq b_{t+l}^{(t+2)} \leq \dots \leq b_{t+l}^{(t+k)},\end{aligned}$$

and $\Psi_{t+1}^{(t+1,t)} - \Psi_{t+1} < 0$ for several consecutive calendar years t . This then leads to a negative sign of the conditionally expected CDR's for several consecutive years. Analogously, this argument applies in the case of a decreasing effect which is first underestimated. Therefore, in the cases where the diagonal effects change in the same direction (increase or decrease) over several years, it is likely that this effect is underestimated and that one observes the same sign of the CDR over the corresponding years.